Unsupervised classification of Mercury’S Visible–Near-Infrared MASCS/MESSENGER reflectance spectra for automated surface mapping.

Mario D’Amore

(mario.damore@dlr.de)

Introduction - Instrument

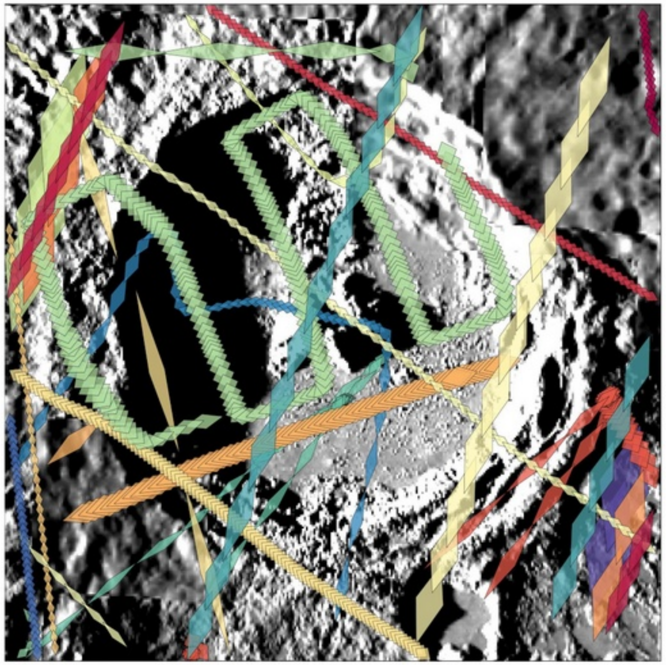
- MESSENGER (MErcury Surface, Space ENvironment, GEochemistry, and Ranging) was a NASA scientific investigation of the planet Mercury.
- Launched in August 2004, Mercury orbit in March 2011 , completed primary mission on March 2012. Impacted the surface of the planet on April 30, 2015 concluding 4 years of extendend mission.
- Collected ~100,000 images and ~5 million spectra in the 320-1450 nm range with the Mercury Atmospheric and Surface and Composition Spectrometer (MASCS) instrument.
- The goal of the prime mission was to understand Mercury, still the least explored terrestrial planet, and how it was formed in order to better understand the other terrestrial planets and their evolution.
- Only Mariner 10 (1975) visited Mercury before MESSENGER and we knew little more than Mercury's average density (the second greatest of all the planets), the composition of its atmosphere (thinnest of the terrestrial planets), the fact that it possesses a global magnetic field, and its extreme variations in temperature.
- The key result from messenger are : Volatile-Rich Planet, Polar Deposits, Offset Magnetic Field, Hollows (Mercury specific surface formation), Volcanic Deposits, Global Contraction, Seasonal Exosphere, Dynamic Magnetosphere, Energetic Electrons, Field-Aligned Currents.

Abstract

- Under the hypothesis that surface compositional information can be effectively derived from spectral reflectance measurements with the use of statistical techniques, we try to identify and characterize spectral units from all orbital observations made with MASCS during the primary mission.
  - ICA for data decompostion / compression.
  - manifold learning (UMAP) to project the data in low dimensional (2D) space.
  - hierarchical agglomerative clustering.
- The results indicate a dichotomy in major units, with one spectral unit in the polar regions and a spectrally distinct unit in equatorial areas.
- The spatial extent of the polar unit in the northern hemisphere shows a generally good correlation with the northern volcanic plains and with the regions of enhanced potassium abundance as mapped by the Gamma-Ray Spectrometer on MESSENGER.
- This asymmetry indicates that peak surface temperature, although potentially important, is not the only factor that contributes to the different spectral character of the surface units. We conclude that compositional differences also contribute to the spectral differences between the two major units on the surface of Mercury identified with this hierarchical clustering approach.

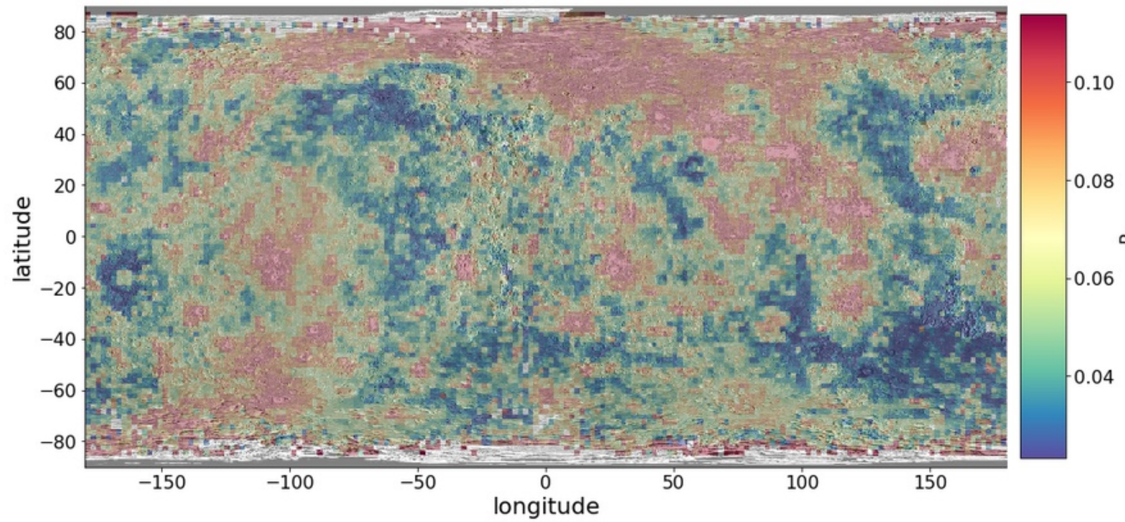
MASCS/MESSENGER Data

1. MASCS VIRS observations in the region of the Kuiper crater, a 62-km-diameter crater with a central peak located at 11.35 S 31.23 W on Mercury.Kuiper shows the highest albedo of the planet and has a fresh ray system, suggesting that it is among the youngest craters of Mercury.

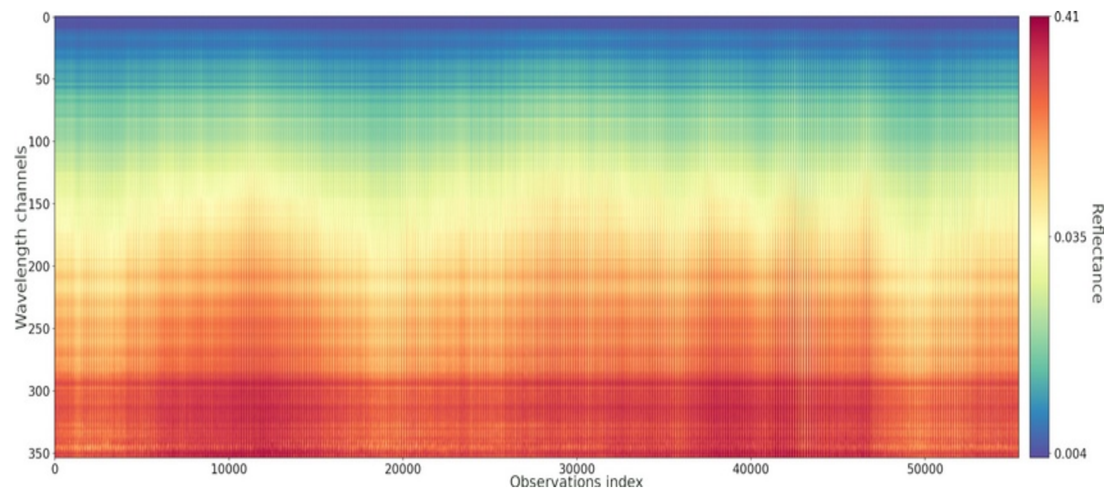


Each polygon represent a approximately a VIS and NIR observation. Note how different orbits (color coded) show different polygons size due to different observation geometry or spacecraft altitude.

2. Square polygons are used as base pixel to construct a global datacube, using the median value of all the observation within as pixel reflectance. Data here are reflectance at 700 nm. Data dimension [55399, 396].



3.Spectrogram of the data in fig **2.**



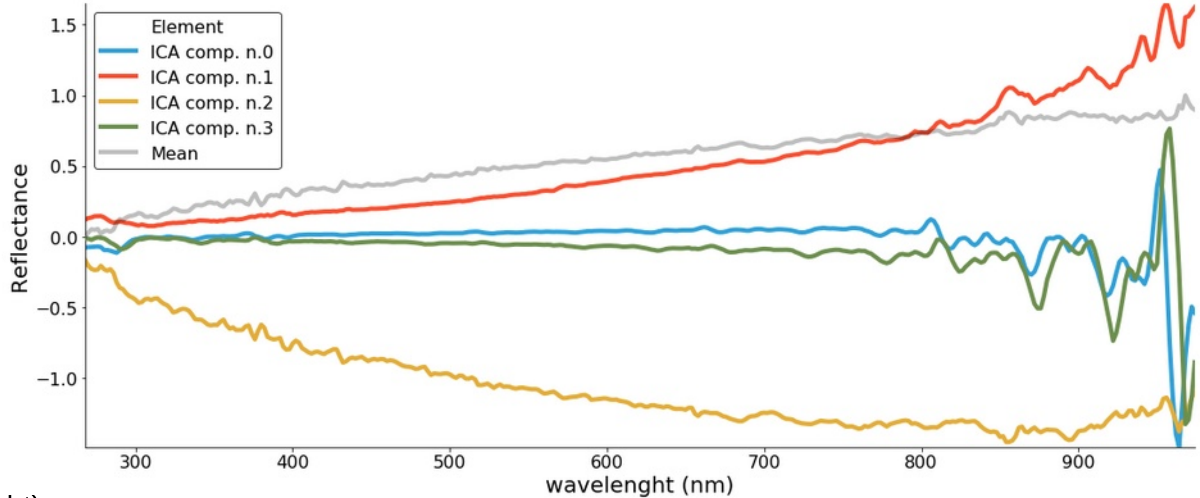
Independent component analysis (ICA)

The data are decomposed via Independent component analysis (ICA), a signal processing technique for separating multivariate signal into additive sub components. Data dimension from [55399,396] to [55399, 4].

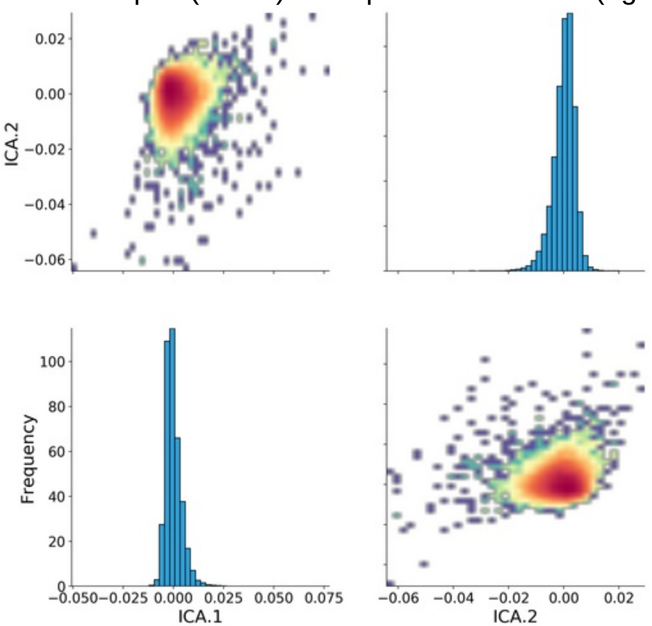
ICA tries to find a linear transformation **W** so **s**=**Wx**.

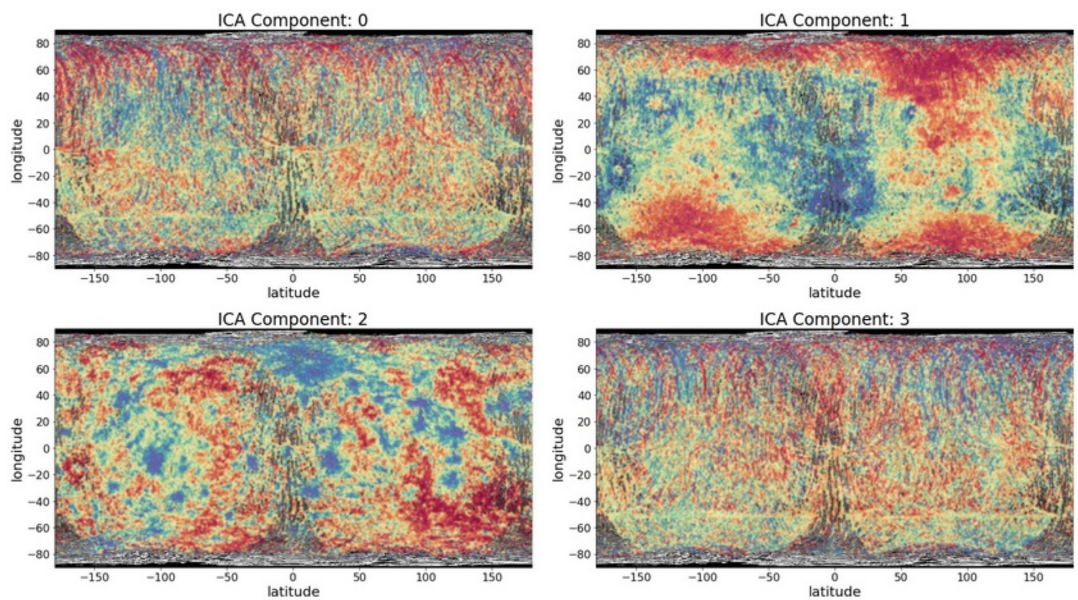
s (fig in the right) is a vector of maximally independent components, **x** are the initial data (previous slide).

W is also called the mixing matrix.



W loading for each component $\mathbf{s}_i$
Scatterplot (below) and spatial distribution (right)



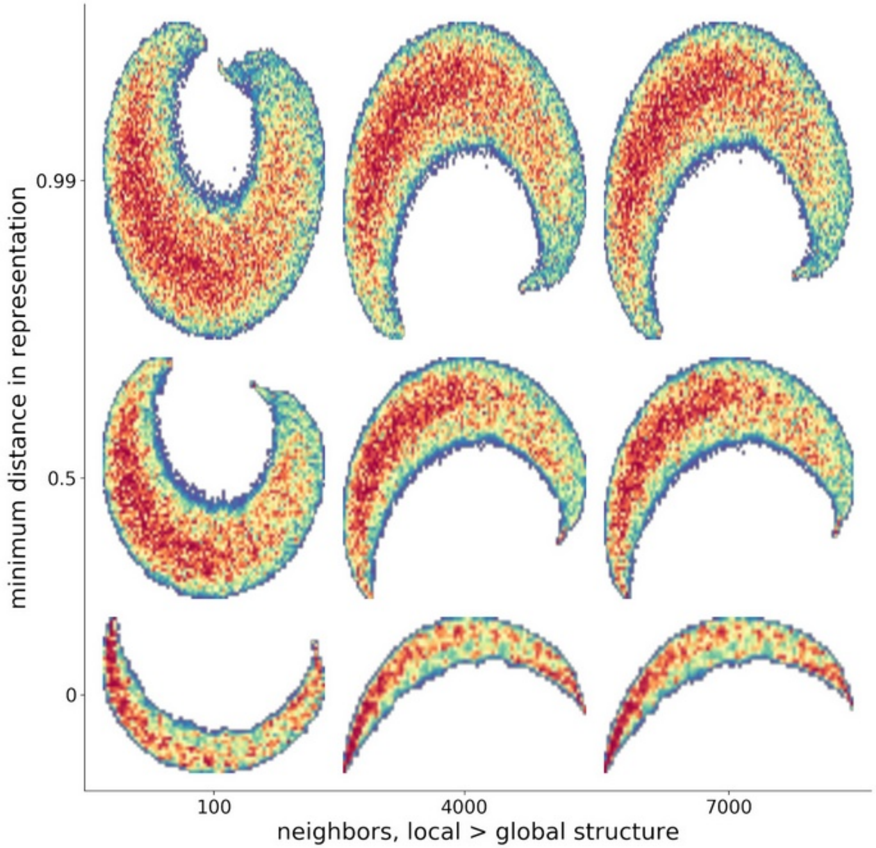


Uniform Manifold Approximation and Projection (UMAP)

UMAP compress data dimension from [55399,4] to [55399, 2].

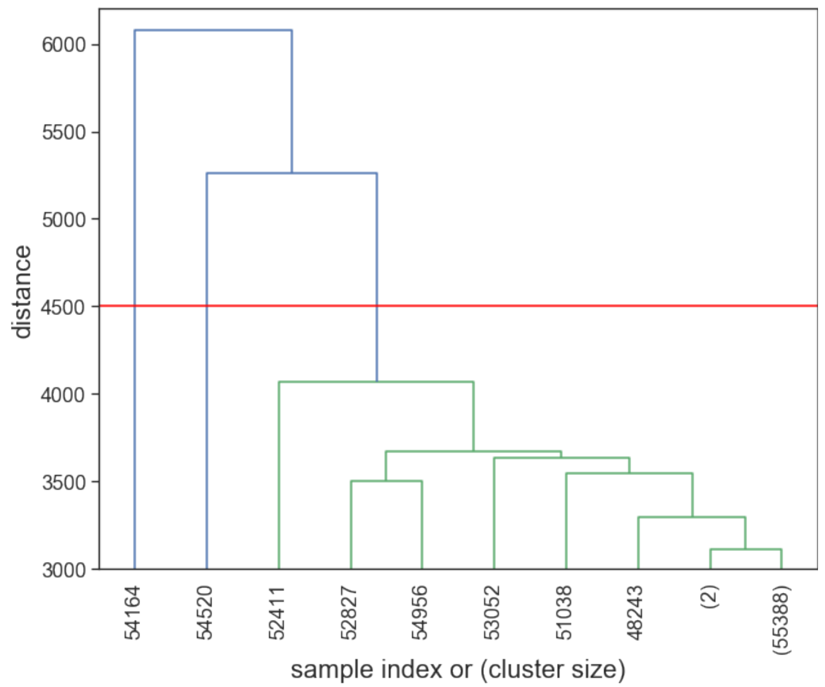
- UMAP uses local manifold approximations, represent those with a local fuzzy simplicial set representations and uses their union to construct a topological representation of the high dimensional data.
- Given some low dimensional representation of the data, a similar process can be used to construct an equivalent topological representation in the other direction (low to higher dimension).
- UMAP then searches for a low dimensional projection of the data that has the closest possible equivalent fuzzy topological structure to the high dimensional one.
- hyperparameters
  - *n_neighbors* : sensitivity to local or global structures in the data by constraining the size of the local neighborhood. Values can go from 0 (local) to the size of the data (global).
  - *min_dist :* minimum distance apart that points are allowed to be in the low dimensional representation. Low values will result in more dense embeddings while larger values will result in more sparse embeddings.

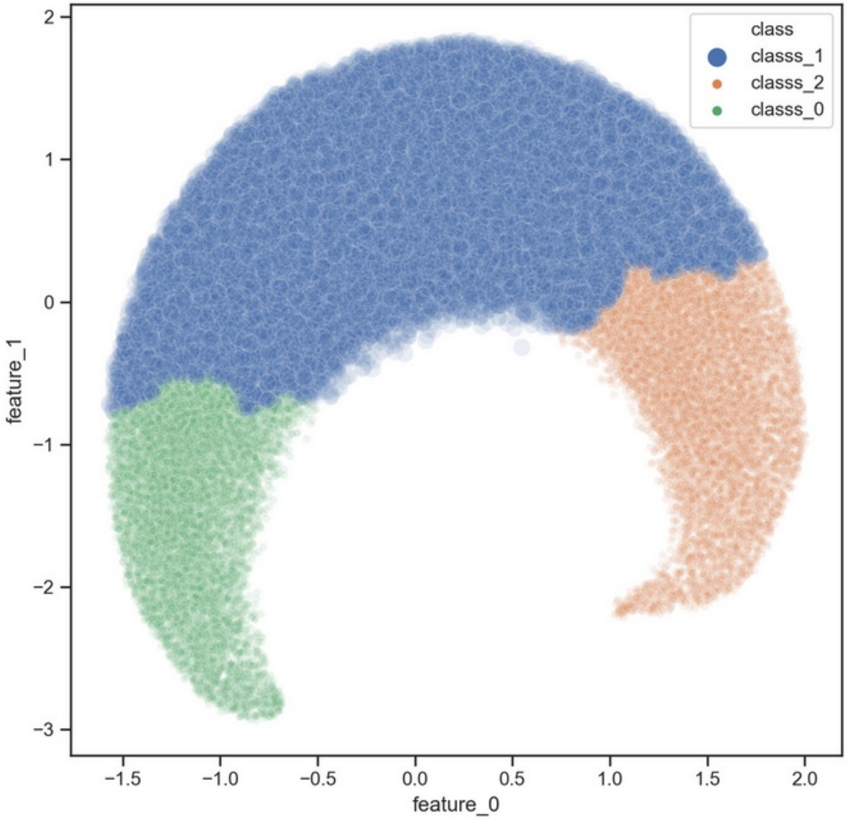
A rigorous mathematical background could be fund in https://umap-learn.readthedocs.io/en/latest/how_umap_works.html



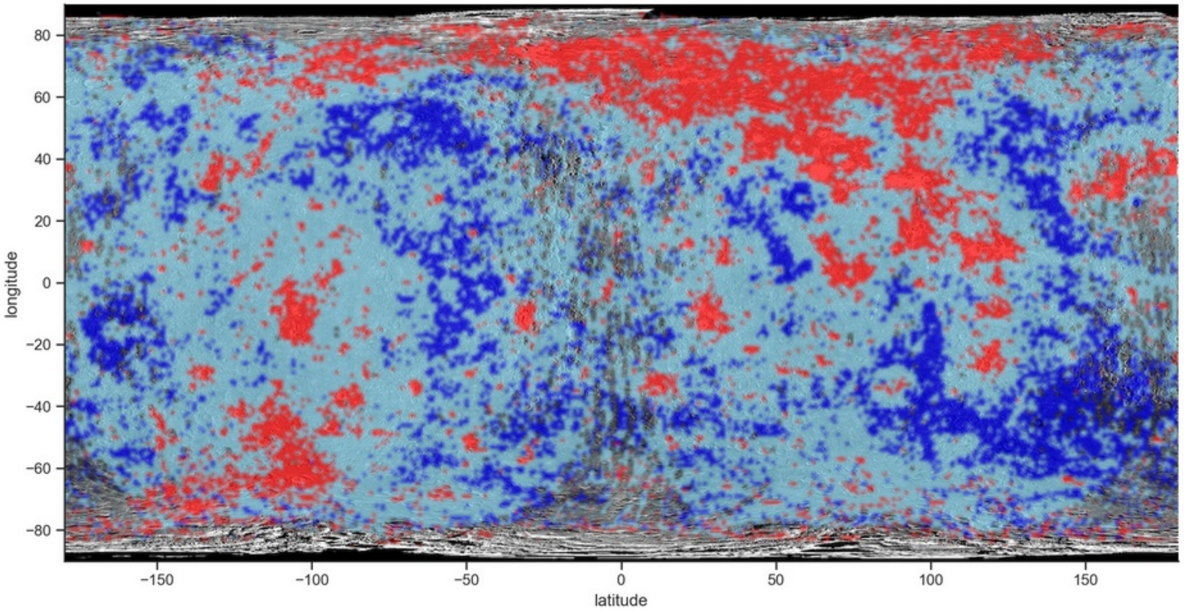
Unsupervised Classification : Hierarchical Clustering

- Hierarchical clustering (HC) algorithms build nested clusters by merging or splitting them successively.
- HC merges closer points together, irrespective of the final cluster class balance.
- The hierarchy of clusters is represented as a dendrogram.
- The final tree has one unique cluster comprising all the samples at the root and clusters with only one sample at the leaves (bottom of the tree).
- The branch lengths represent the distance between the child clusters.
- Small gaps connect more similar clusters and big gaps connect more different clusters.
- The cluster distance is computed as the maximum Euclidean distance distances between all observations of the two clusters.

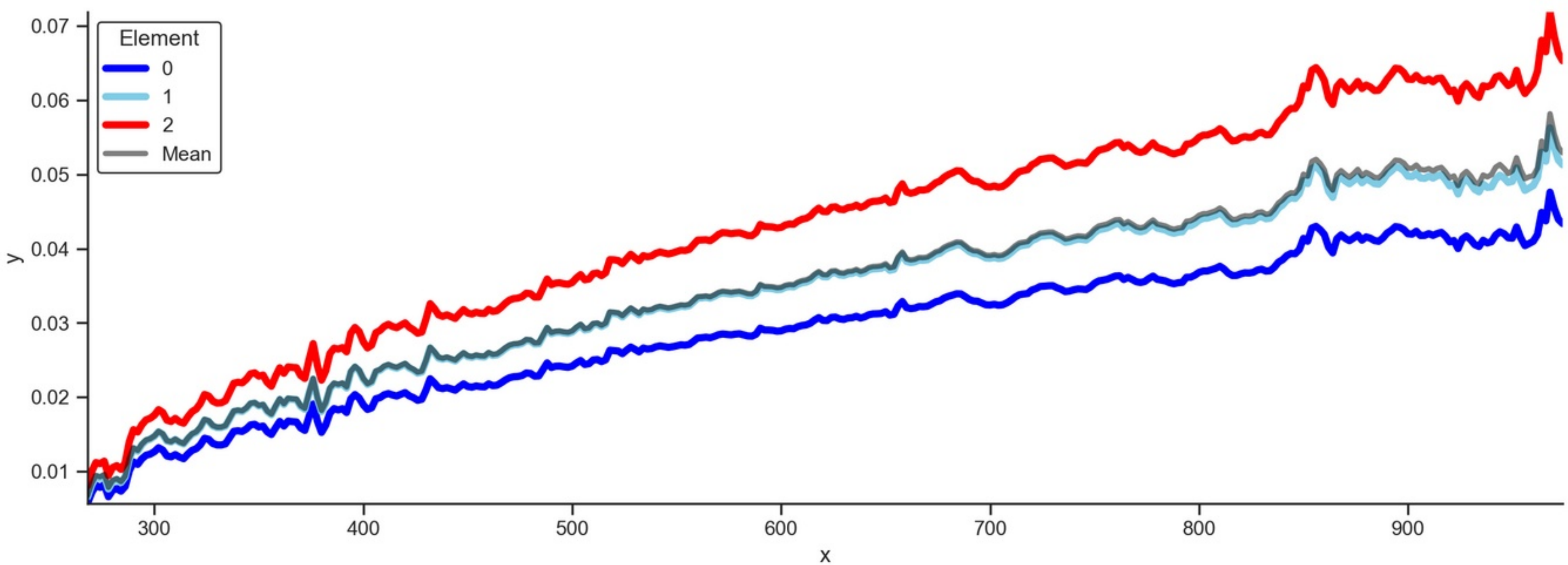




Unsupervised Classification : Hierarchical Clustering

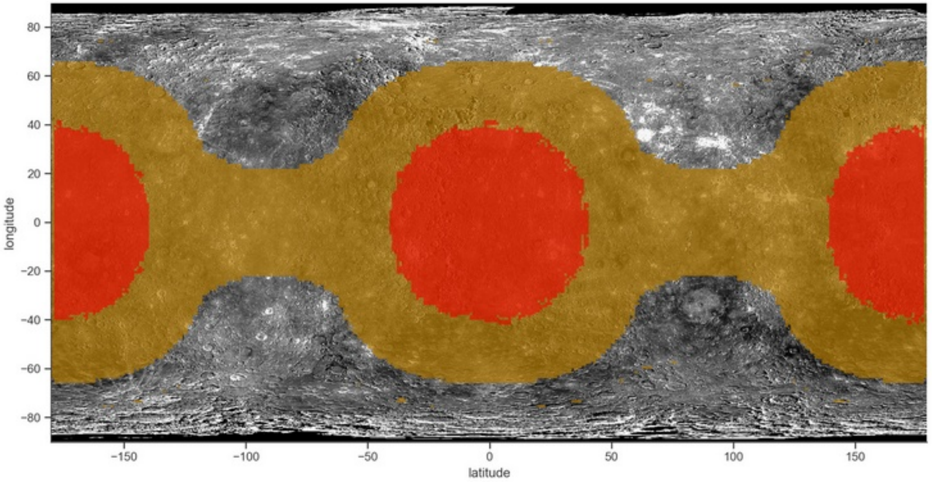


Surface unsupervised classification spatial distribution (**left**) and average classes speactra (**below**)

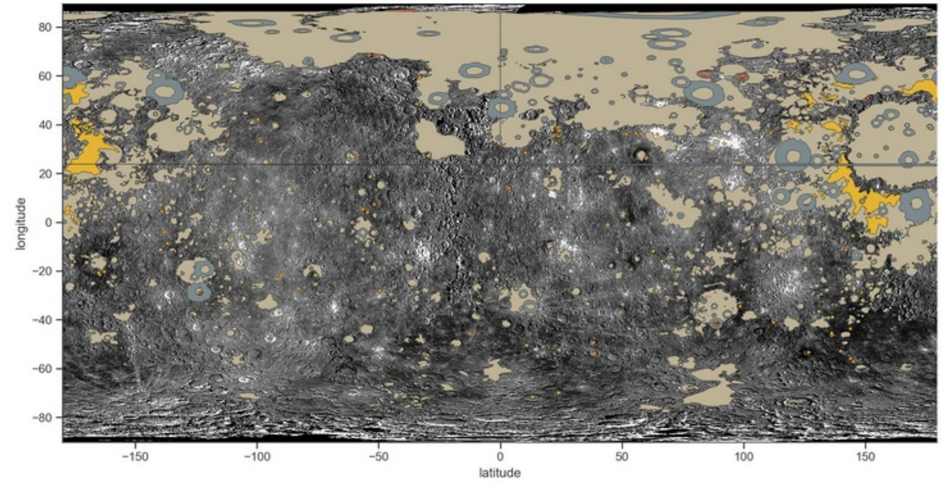


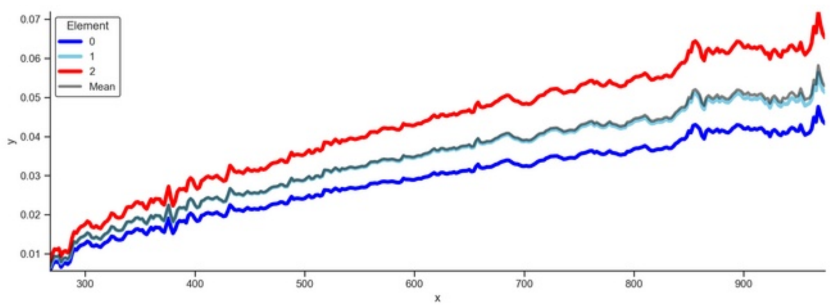
Comparison

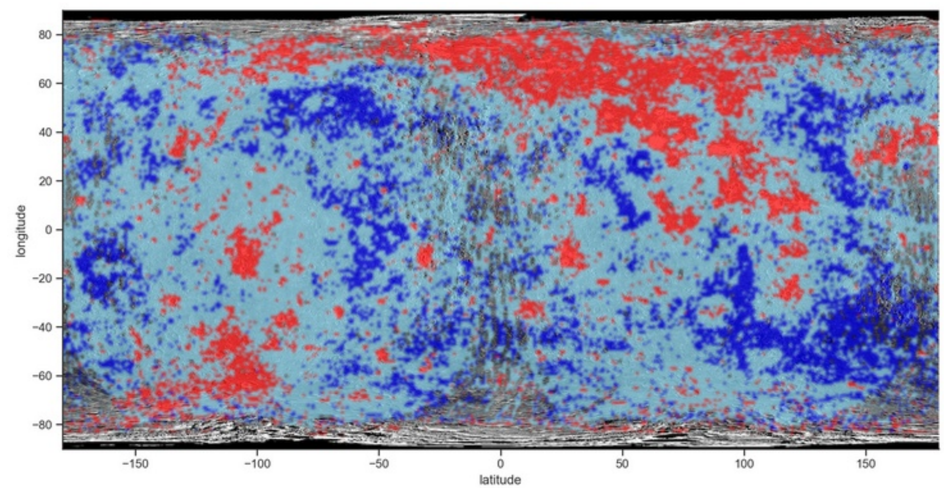
Maximum surface temperature: red > 690 K, brown > 550K (Vasavada, 1999)



Mercury smooth plains (Denevi et al., 2013)







Surface unsupervised classification spatial distribution (**above**) and average classes speactra (**right**)

Denevi, B. et al, The distribution and origin of smooth plains on Mercury. Journal of Geophysical Research: Planets, 118(5):891–907, May 2013, doi: 10.1002/jgre.20075.
Vasavada, a. Near-Surface Temperatures on Mercury and the Moon and the Stability of Polar Ice Deposits. Icarus, 141(2):179–193, October 1999, doi: 10/b9fhjd

BACKUP SLIDES

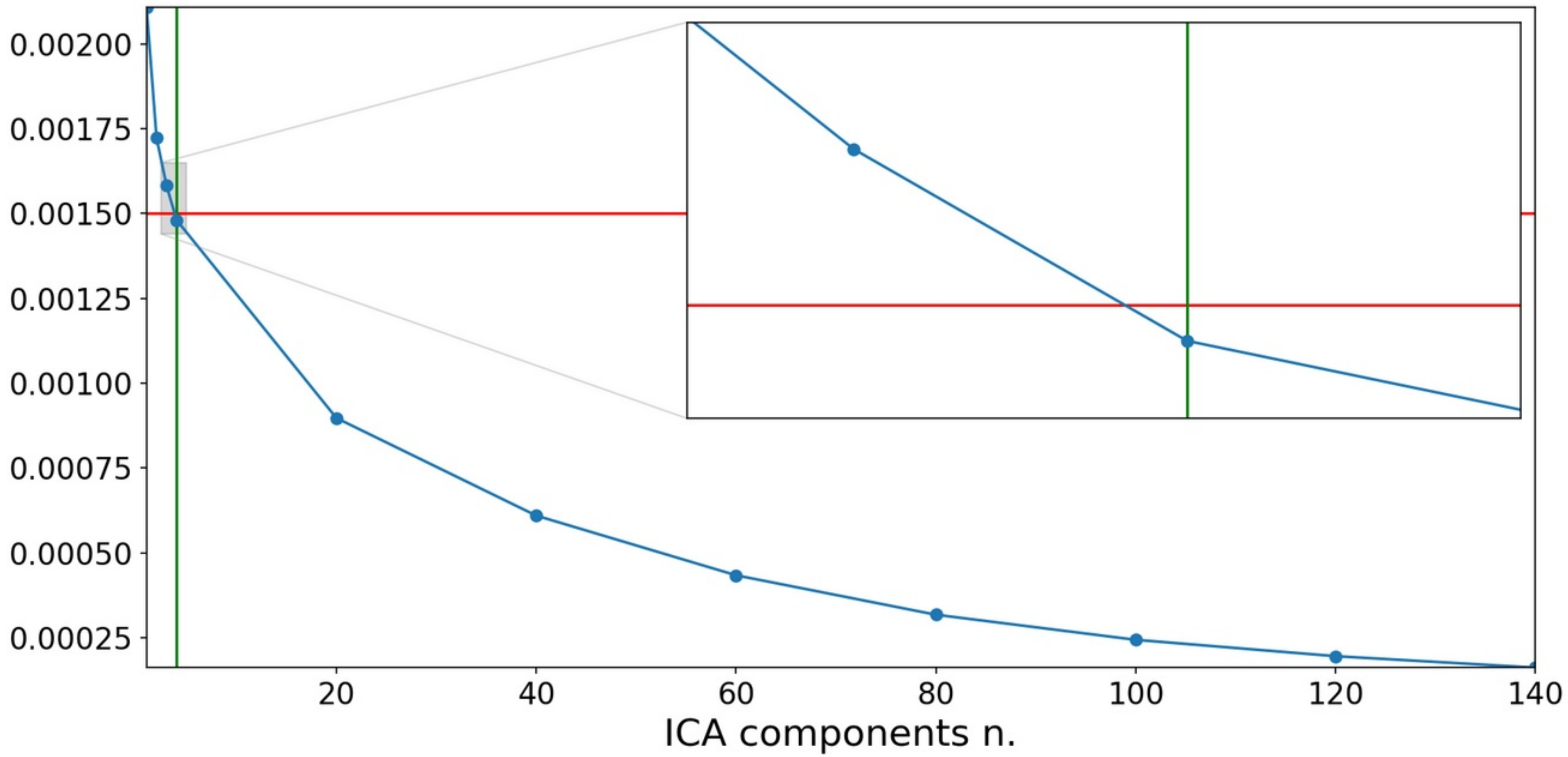
Independent component analysis (ICA)

Determining the minimum number of components for accurate representation.

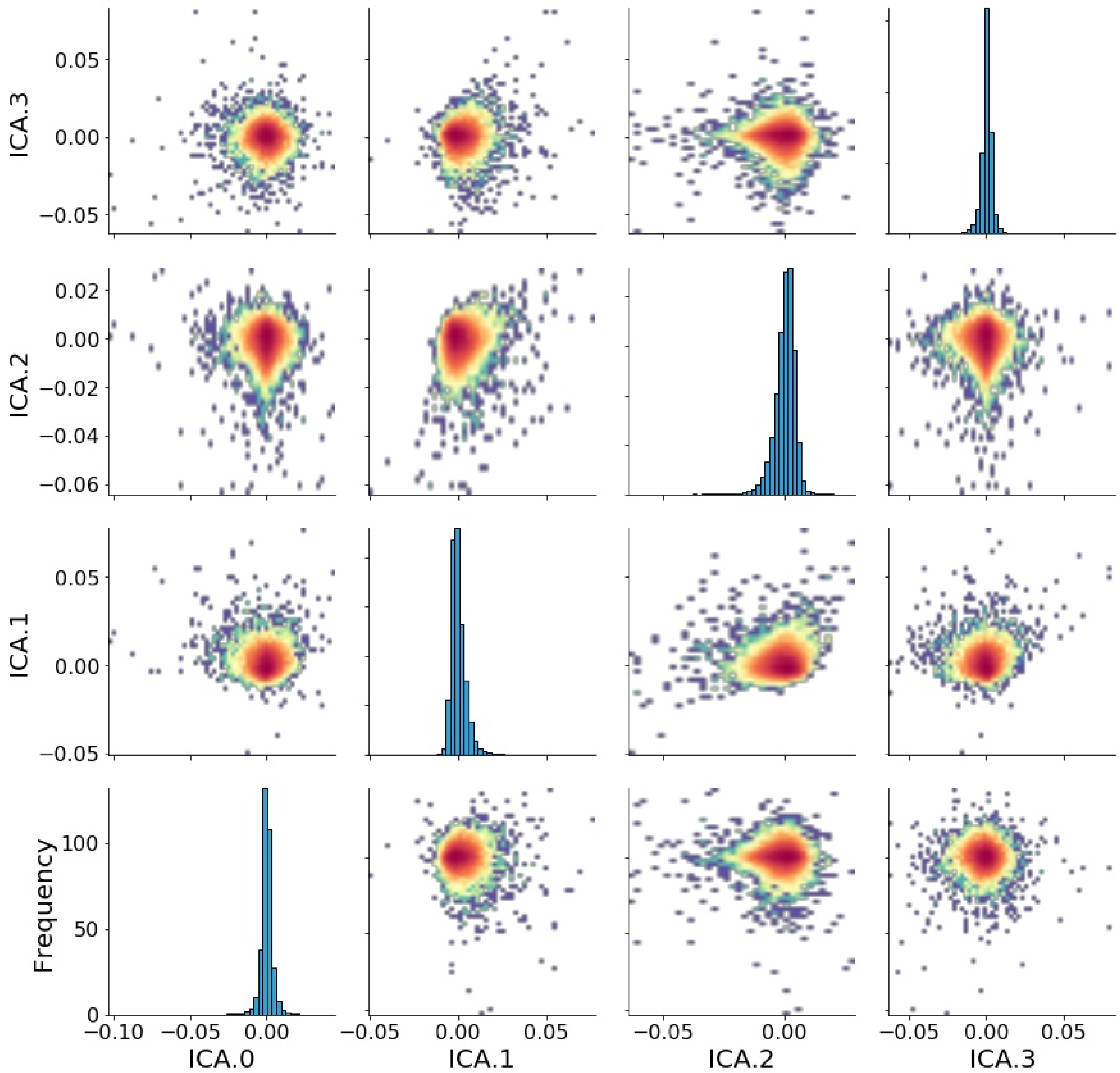
Increasing the number of k of independent components **s** the reconstruction error decreases.

When reconstruction error $||x-s(k)||<0.0015$ we reached the threshold that represents the noise in the data.

K =3



Independent component analysis (ICA)



Hierarchical Clustering vs K-Means